

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Lal, Preeti  
Corley, Neil C.
- (ii) TITLE OF THE INVENTION: HUMAN SHORT CHAIN DEHYDROGENASE
- (iii) NUMBER OF SEQUENCES: 3
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
  - (B) STREET: 3174 Porter Dr.
  - (C) CITY: Palo Alto
  - (D) STATE: CA
  - (E) COUNTRY: USA
  - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Diskette
  - (B) COMPUTER: IBM Compatible
  - (C) OPERATING SYSTEM: DOS
  - (D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: To Be Assigned
  - (B) FILING DATE: Filed Herewith
  - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
  - (A) APPLICATION NUMBER:
  - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Billings, Lucy J.
  - (B) REGISTRATION NUMBER: 36,749
  - (C) REFERENCE/DOCKET NUMBER: PF-0475 US
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 650-855-0555
  - (B) TELEFAX: 650-845-4166
  - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 313 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
  - (A) LIBRARY: PROSNOT01
  - (B) CLONE: 356351
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

Met Ala Ala Pro Met Asn Gly Gln Val Cys Val Val Thr Gly Ala Ser
 1           5           10           15
Arg Gly Ile Gly Arg Gly Ile Ala Leu Gln Leu Cys Lys Ala Gly Ala
      20           25           30
Thr Val Tyr Ile Thr Gly Arg His Leu Asp Thr Leu Arg Val Val Ala
      35           40           45
Gln Glu Ala Gln Ser Leu Gly Gly Gln Cys Val Pro Val Val Cys Asp
      50           55           60
Ser Ser Gln Glu Ser Glu Val Arg Thr Leu Phe Glu Gln Val Asp Arg
      65           70           75           80
Glu Gln Gln Gly Arg Leu Asp Val Leu Val Asn Asn Ala Tyr Ala Gly
      85           90           95
Val Gln Thr Ile Leu Asn Thr Arg Asn Lys Ala Phe Trp Glu Thr Pro
      100          105          110
Ala Ser Met Trp Asp Asp Ile Asn Asn Val Gly Leu Arg Gly His Tyr
      115          120          125
Phe Cys Ser Val Tyr Gly Ala Arg Leu Met Val Pro Ala Gly Gln Gly
      130          135          140
Leu Ile Val Val Ile Ser Ser Pro Gly Ser Leu Gln Tyr Met Phe Asn
      145          150          155          160
Val Pro Tyr Gly Val Gly Lys Ala Ala Cys Asp Lys Leu Ala Ala Asp
      165          170          175
Cys Ala His Glu Leu Arg Arg His Gly Val Ser Cys Val Ser Leu Trp
      180          185          190
Pro Gly Ile Val Gln Thr Glu Leu Leu Lys Glu His Met Ala Lys Glu
      195          200          205
Glu Val Leu Gln Asp Pro Val Leu Lys Gln Phe Lys Ser Ala Phe Ser
      210          215          220
Ser Ala Glu Thr Thr Glu Leu Ser Gly Lys Cys Val Val Ala Leu Ala
      225          230          235          240
Thr Asp Pro Asn Ile Leu Ser Leu Ser Gly Lys Val Leu Pro Ser Cys
      245          250          255
Asp Leu Ala Arg Arg Tyr Gly Leu Arg Asp Val Asp Gly Arg Pro Val
      260          265          270
Gln Asp Tyr Leu Ser Leu Ser Ser Val Leu Ser His Val Ser Gly Leu
      275          280          285
Gly Trp Leu Ala Ser Tyr Leu Pro Ser Phe Leu Arg Val Pro Lys Trp
      290          295          300
Ile Ile Ala Leu Tyr Thr Ser Lys Phe
305          310

```

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1387 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PROSNOT01
- (B) CLONE: 356351

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

CTAACTTTGG CCTGGGACTC TGCCCTCTA CCTCAGCACA GAATCGCCCC GGGTCCTACT      60
ACAGAATCAA TCCTTGAACA CTGCCTCCAC GTCGCCGGCT CAATCTGGGC GAGAACCCAG      120
ACTTCCACCG CAGCCCCGCA ATCTGCAGAC CTCAGCGGCA GCGCAGGTGG CAGACCTGCC      180
TCCTTTGCCT GTGAGTCATG GCAGCTCCCA TGAATGGCCA AGTGTGTGTG GTGACTGGTG      240

```

```

CCTCCAGGGG TATTGGCCGT GGCATTGCCT TGCAGCTCTG CAAAGCAGGC GCCACAGTTT 300
ACATCACTGG CCGCCATCTG GACACCCCTC GCGTTGTTGC TCAGGAGGCA CAATCCCTCG 360
GGGGCCAATG TGTGCCTGTG GTGTGCGATT CAAGCCAGGA GAGTGAAGTG CGAACGCTGT 420
TTGAGCAAGT GGATCGGGAA CAGCAAGGGC GTCTAGATGT GCTGGTCAAC AATGCTTATG 480
CAGGGGTCCA GACGATCCTG AACACCAGGA ATAAGGCATT CTGGGAAACC CCTGCCTCCA 540
TGTGGGATGA TATCAACAAC GTCGGACTCA GAGGCCACTA CTTTGTCTCA GTGTATGGGG 600
CACGGCTGAT GGTACCAGCT GGCCAGGGGC TCATCGTGTT CATCTCCTCC CCAGGAAGCC 660
TGCAGTATAT GTTCAATGTC CCCTATGGTG TGGGCAAAGC TGCCTGTGAC AAGCTGGCTG 720
CTGACTGTGC CCACGAGCTG CGGCGCCATG GGGCTAGCTG TGTGTCTCTG TGGCCGGGGA 780
TTGTGCAGAC AGAAGTCTGT AAGGAGCATA TGGCAAAGGA GGAGTCTCTG CAGGATCCTG 840
TGTTGAAGCA GTTCAAATCA GCCTTCTCAT CTGCAGAAAC CACAGAATTG AGTGGCAAAT 900
GTGTGGTGCG TTTGGCAACA GATCCCAATA TCCTGAGCCT GAGTGGTAAG GTGCTGCCAT 960
CCTGTGACCT TGCTCGACGC TATGGCCCTC GGGATGTGGA CGGCCGCCCC GTCCAAGACT 1020
ATTTGTCTTT GAGCTCTGTT CTCTCACACG TGTCCGGCCT GGGCTGGCTG GCCTCCTACC 1080
TGCCCTCCTT CCTCCGTGTG CCCAAGTGGA TTATTGCCCT CTACACTAGC AAGTTCTAAC 1140
CCTCCTGGTC TGACACTACG TCTCTGCTTG TCTTCTCATT TGGACTTGGT GGTTCGTCCT 1200
GTCTCAGTGA AACAGCAGCC TTTCTTGTGT ACCCATACCC TTGATATGAA GAGAAGCCCT 1260
CTGCTGTGTG TCCGTGGTGA GTTCTGGGGT GCGCCTAGGT CCCTTCTTTG TGCCTTGTTT 1320
TTCCTTGTC TTTCTTTTAC TTTTTCGCTT AGTATTGAAA AATGCTCTTG GAGCTAATAA 1380
AAGTCTA 1387

```

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 323 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 2315796

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Gly Val Ile Leu Gln Asp Gln Val Ala Leu Val Thr Gly Ala Ser
 1           5           10           15
Arg Gly Ile Gly Arg Gly Ile Ala Leu Gln Leu Gly Glu Ala Gly Ala
 20           25           30
Thr Val Tyr Ile Thr Gly Arg Arg Pro Glu Leu Ser Asp Asn Phe Arg
 35           40           45
Leu Gly Leu Pro Ser Leu Asp Tyr Val Ala Lys Glu Ile Thr Ser Arg
 50           55           60
Gly Gly Lys Gly Ile Ala Leu Tyr Val Asp His Ser Asn Met Thr Glu
 65           70           75           80
Val Lys Phe Leu Phe Glu Lys Ile Lys Glu Asp Glu Glu Gly Lys Leu
 85           90           95
Asp Ile Leu Val Asn Asn Val Tyr Asn Ser Leu Gly Lys Ala Thr Glu
100           105           110
Met Ile Gly Lys Thr Phe Phe Asp Gln Asp Pro Ser Phe Trp Asp Asp
115           120           125
Ile Asn Gly Val Gly Leu Arg Asn His Tyr Tyr Cys Ser Val Tyr Ala
130           135           140
Ala Arg Met Met Val Glu Arg Arg Lys Gly Leu Ile Val Asn Val Gly
145           150           155           160
Ser Leu Gly Gly Leu Lys Tyr Val Phe Asn Val Ala Tyr Gly Ala Gly
165           170           175
Lys Glu Ala Leu Ala Arg Met Ser Thr Asp Met Ala Val Glu Leu Asn
180           185           190

```

PF-0475-2 DIV

Pro	Tyr	Asn	Val	Cys	Val	Val	Thr	Leu	Ile	Pro	Gly	Pro	Val	Lys	Thr
		195					200				205				
Glu	Thr	Ala	Asn	Arg	Thr	Ile	Ile	Asp	Asp	Ala	Tyr	Lys	Met	Ile	Lys
		210				215				220					
Glu	Asn	Pro	Glu	Leu	Glu	Glu	Phe	Ile	Lys	Gly	Glu	Ser	Thr	Glu	Tyr
225					230					235				240	
Thr	Gly	Lys	Ala	Leu	Ala	Arg	Leu	Ala	Met	Asp	Pro	Gly	Lys	Leu	Lys
			245						250					255	
Lys	Ser	Gly	Lys	Thr	Leu	Phe	Thr	Glu	Asp	Leu	Ala	Gln	Lys	Tyr	Asp
			260					265					270		
Phe	Ser	Asp	Lys	His	Gly	Ala	Gly	Met	Glu	Pro	Gln	Asn	Ile	Arg	Ser
		275					280					285			
Ile	Arg	Thr	Ile	Leu	Gly	Thr	Met	Gly	Lys	Glu	Glu	Val	Ala	Lys	Tyr
		290				295					300				
Ile	Pro	Pro	Gln	Ile	Lys	Leu	Pro	Lys	Trp	Val	Ile	Trp	Gln	Ser	Val
305					310					315					320
Asn	Arg	Phe													